The EMBO Journal Vol.17 No.23 pp.6819–6826, 1998

Crystal structure of methionyl-tRNA$_f^{Met}$ transformylase complexed with the initiator formyl-methionyl-tRNA$_f^{Met}$

Emmanuelle Schmitt, Michel Panvert, Sylvain Blanquet and Yves Mechulam[1]

Laboratoire de Biochimie, Unité Mixte de Recherche No. 7654 du Centre National de la Recherche Scientifique, Ecole Polytechnique, F-91128 Palaiseau cedex, France

[1]Corresponding author
e-mail: yves@botrytis.polytechnique.fr

The crystal structure of *Escherichia coli* methionyl-tRNA$_f^{Met}$ transformylase complexed with formyl-methionyl-tRNA$_f^{Met}$ was solved at 2.8 Å resolution. The formylation reaction catalyzed by this enzyme irreversibly commits methionyl-tRNA$_f^{Met}$ to initiation of translation in eubacteria. In the three-dimensional model, the methionyl-tRNA$_f^{Met}$ formyltransferase fills in the inside of the L-shaped tRNA molecule on the D-stem side. The anticodon stem and loop are away from the protein. An enzyme loop is wedged in the major groove of the acceptor helix. As a result, the C1–A72 mismatch characteristic of the initiator tRNA is split and the 3′ arm bends inside the active centre. This recognition mechanism is markedly distinct from that of elongation factor Tu, which binds the acceptor arm of aminoacylated elongator tRNAs on the T-stem side.

Keywords: crystalline structure/formylation/transfer RNA/translation initiation

Introduction

Transfer RNAs play a central role in protein biosynthesis by providing the link between the genetic message and functional proteins. To achieve their function, tRNAs must first undergo aminoacylation. The past few years have seen much progress in the description of the specific recognition of tRNAs by their cognate aminoacyl-tRNA synthetases.

Once esterified, tRNAs are channeled towards the appropriate site on the ribosome. Two sites can be distinguished: the A-site, where all elongator tRNAs converge, and the P-site, in which a specific machinery drives initiator tRNA. Initiator tRNA has sequence features distinguishing it from elongator tRNAs. These features have been shown to support the specialized function of this tRNA, including a susceptibility to formylation in eubacteria. Commitment of the initiator tRNA to the P-site through the recruitment of initiator factor IF2 follows this modification catalyzed by methionyl-tRNA$_f^{Met}$ formyltransferase (formylase). Formylation also prevents complexation by EF-Tu.GTP, thereby making easier the selection by IF2 (reviewed in Schmitt *et al.*, 1996b). The key role of the formylation step in the channeling of the initiator tRNA towards the ribosomal P-site is underlined by the observation that inactivation of the formylase gene on the *Escherichia coli* chromosome severely impairs cell growth (Guillon *et al.*, 1992a).

The main basis for the specific formylation of eubacterial methionyl-tRNA$_f^{Met}$ is the lack of base pairing at the top of the acceptor helix. In the *E.coli* initiator tRNA, this defect is a C1–A72 mismatch. The A73 discriminator base and base pairs G2–C71, C3–G70 and G4–C69 in the acceptor arm also contribute to the specification of the formylation reaction (Lee *et al.*, 1991; Guillon *et al.*, 1992b). The present work introduces a three-dimensional (3D) model of the crystal structure of *E.coli* formylase complexed with its formyl-methionyl-tRNA$_f^{Met}$ product. Comparison of this model with that of native formylase (Schmitt *et al.*, 1996a), solved at 2.8 and 2.0 Å resolution, respectively, enables us to describe the relation between tRNA$_f^{Met}$ structure and its unique property to undergo formylation beyond esterification by a methionine.

Results and discussion

Structure determination

The structure of formylase complexed with formyl-methionyl-tRNA$_f^{Met}$ was determined. This ligand, which firmly binds the formylase (Kahn *et al.*, 1980), was chosen instead of the methionyl-tRNA$_f^{Met}$ substrate because it displays much higher chemical stability in solution. Structure solution was performed by single isomorphous replacement with the help of molecular replacement using the native enzyme structure. The asymmetric unit consists of two complexed proteins related by a nearly exact translation of a half-cell edge along the *b* axis. One complex is, on the whole, slightly more disordered than the other. However, neither the free-*R* factor nor the quality of the density could be improved by refining the structure without NCS (non-crystallographic symmetry) restraints. Therefore, all refinement was performed with NCS-restraints excluding regions involved in crystal packing. The final crystallographic R and free *R*-factors are 24.7 and 29.2%, respectively. All enzyme and tRNA residues could be constructed and refined with good geometry. The final electron density is generally of good quality (Figure 1). However, in the tRNA molecules, the D- and T-loops (residues 16–19), as well as bases 1 and 37, appeared more disordered, even though the positions of these bases could be tentatively modeled.

Overall structure

Native formylase is composed of two domains connected by an elongated linker (Schmitt *et al.*, 1996a). The catalytic N-terminal domain (residues 1–189) contains a Rossmann fold and shares a high level of similarity with *E.coli* glycinamide ribonucleotide transformylase (GARF), which also uses 10-formyltetrahydrofolate (FTHF) as the

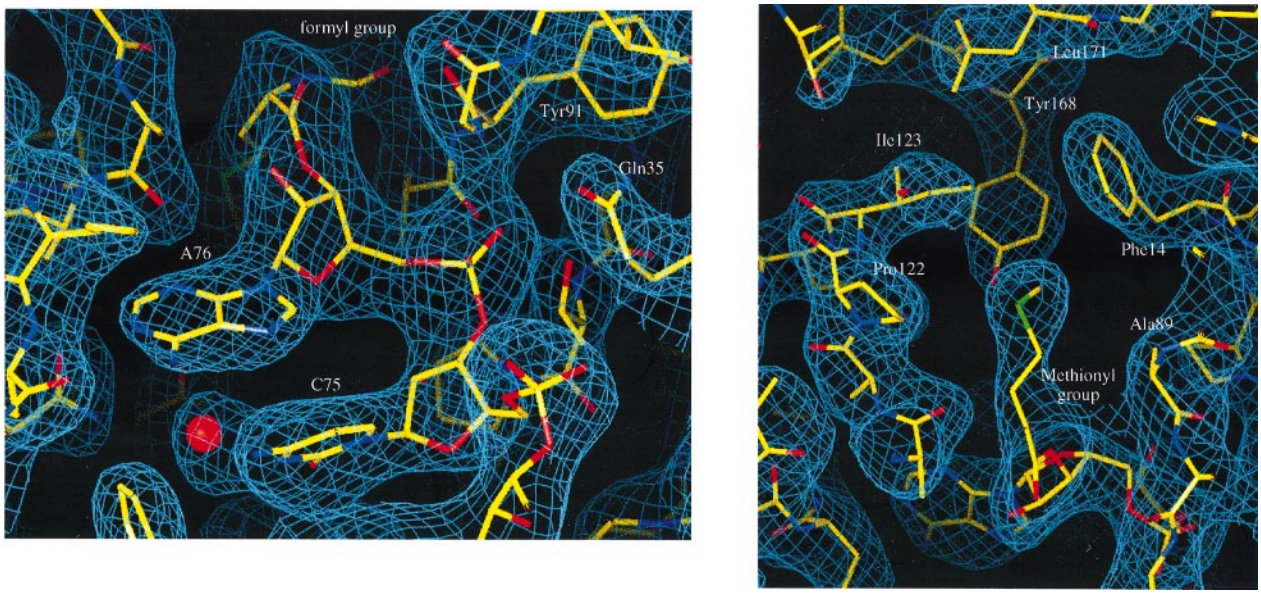

Fig. 1. Views of the 2.8 Å resolution '$2F_o–F_c$' map contoured at 1.0 standard deviation. The left panel shows bases C75 and A76 with the esterified formyl-methionyl group, as well as the entrance of the catalytic crevice. The strong extra density facing O2 of C75 and stacked onto the ring of A76 is indicated by a red sphere. Note that Tyr90 is incorrectly labelled as 'Tyr91'. The right panel illustrates the hydrophobic cavity filled by the methionyl group esterified to the tRNA.

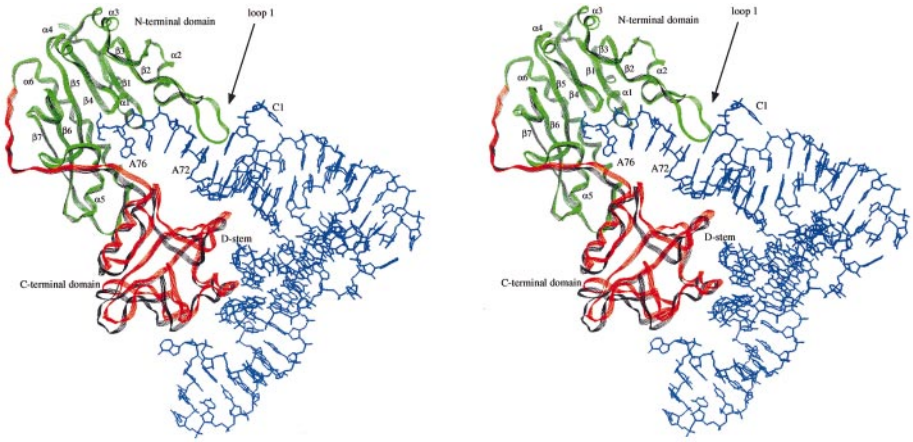

Fig. 2. Stereo representation of the formylase–fMet-$tRNA_f^{Met}$ complex. The tRNA molecule is represented by blue bonds. Positions of the D-stem as well as of bases A76, A72 and C1 are indicated. The secondary structure elements in the N-domain are indicated and numbered as in Schmitt *et al.* (1996a). Note that although base C1 is clearly oriented outside the acceptor stem, its precise positioning is tentative. The complexed formylase is represented as ribbon, with its N-terminal domain in green and the linker peptide and C-terminal domains in red. The structure of the free enzyme (PDB entry 1fmtA) is also shown as a black ribbon, with the N-terminal domains of the free and of the complexed enzyme superimposed. This superimposition emphasizes the movement of the C-terminal domain accompanying the binding of the tRNA molecule. Indeed, upon superimposition of the N-terminal domains, r.m.s. deviations are 0.5 Å for the C_α of the N-terminal domains and 2.6 Å for the C_α of the C-terminal domains. In constrast, the C-terminal domains alone can be superimposed with an r.m.s. deviation of 0.6 Å. The figure was drawn using Setor (Evans, 1993).

formyl donor (Almassy *et al.*, 1992; Chen *et al.*, 1992; Schmitt *et al.*, 1996a). This resemblance enabled description of the active centre of the formylase (Schmitt *et al.*, 1996a). Two idiosyncratic features distinguish formylase from GARF: an additional loop (loop 1, residues 34–49) inside the catalytic domain and a β-barrel C-terminal domain (residues 209–314).

In the complex, both domains of formylase as well as the linker have contacts inside the L-shaped tRNA molecule. As shown in Figure 2, the acceptor arm of tRNA is clamped between one side of the C-terminal domain of the protein and the large loop 1. The β-barrel is bound to both the D-arm and to the minor groove of the acceptor stem. The elongated linker peptide is in contact with the acceptor stem, while the 3′ extremity of tRNA carrying the formylated methionyl group enters the N-terminal domain. Neither the TψC stem nor the D or T loops are involved in the complex formation. Anticodon stem and loop are outside, with the anticodon bases stacked like those of free $tRNA^{Phe}$ (Suddath *et al.*, 1974; Ladner *et al.*, 1975). The three G–C base pairs of the anticodon stem, crucial for the initiator tRNA function,

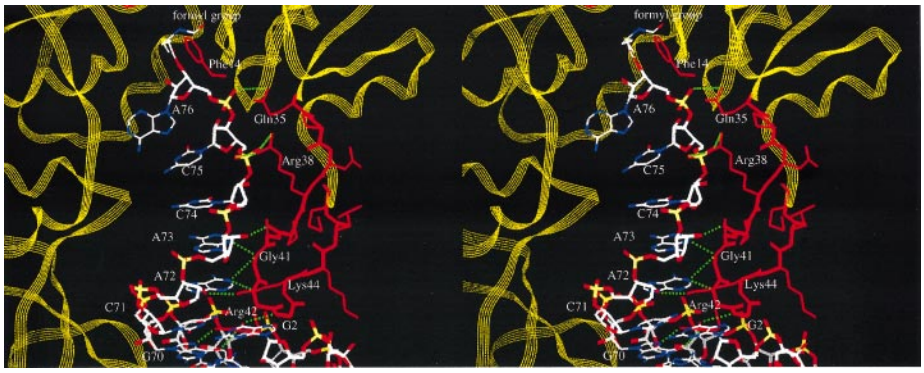

Fig. 3. Close-up stereo view showing the guiding of the top of the tRNA acceptor stem by loop 1 of the formylase. The enzyme is drawn as a yellow ribbon with the exception of loop 1, which is represented as red bonds. The acceptor end of the bound tRNA is represented with solid bonds. Hydrogen bonds are schematized as green dashed lines. The figure was drawn by using Setor (Evans, 1993).

as well as the anticodon loop, do not display any peculiar arrangement. This finding supports the idea that the G–C base pairs rather than a special conformation of the anticodon region are recognized by the components of the ribosomal P-site (Mandal *et al.*, 1996). In the crystal, packing interactions occur between the two T-loops and between the two anticodon loops of adjacent tRNA molecules, while packing of enzyme molecules mainly involves the N-domains. Notably, anticodon loops interact through pairing between $C_{34}A_{35}U_{36}A_{38}$ and the four same bases of an NCS-related tRNA molecule, thereby forming two Watson–Crick A–U pairs and two non-standard C–A pairs.

Superimposition of the peptidic backbones of the N-terminal domains of free and complexed enzymes (Figure 2) reveals a locking movement of the C-domain. The pivoting originates from the Lys206 residue and propagates towards the C-domain. Comparison of the two NCS-related molecules in the native enzyme crystals had already indicated flexibility between the N- and C-domains (Schmitt *et al.*, 1996a). In the complex, the relative movement of the C-domain ensures the filling in of the inside of the L-shape of the polynucleotide. According to the induced fit mechanism theory, this movement is likely to enable the recovering of maximal free energy upon complex formation.

Binding of the initiator tRNA

One striking feature is the opening of the C1–A72 pair in the formylase:tRNA complex, while these bases are face-to-face in free $tRNA_f^{Met}$ (Woo *et al.*, 1980). This difference can be accounted for by the action of the idiosyncratic loop 1, the top of which overhangs the 2–71 pair, and rejects C1 outside of the stem in the solvent. Two sets of contacts between loop 1 and the acceptor stem can be distinguished (Figures 3 and 4). (i) The polypeptide chain runs adjacent to the 3′ strand of the tRNA between bases 72–76. The main chain NH groups of Gly41 and Arg42, and the Nζ of Lys44, interact with N1 and O2′ of A72, respectively. The discriminator base A73 is held through its N3 and O2′ by the main chain NH group of Gly41 and by the carbonyl group of Pro39. The phosphates of C75 and A76 are bound to the side chains of Arg38 and Gln35, respectively, the positions of which are broadly shifted compared with the native enzyme. (ii) The Arg 42 side-chain wedges the major groove of the acceptor stem.

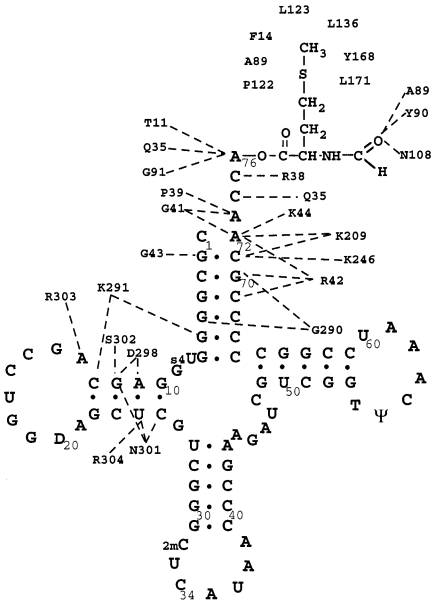

Fig. 4. Main points of contact between formylase and fMet-$tRNA_f^{Met}$. fMet-$tRNA_f^{Met}$ is schematized in the cloverleaf representation. Dotted lines symbolize contacts between the indicated enzyme and tRNA residues. These contacts involve either main chain or side chain atoms (see text). Enzyme residues surrounding the hydrophobic cavity filled by the methionine side chain are also indicated.

It can make two specific hydrogen bonds with the O6 and N7 groups of G70, and its guanidinium group stacks against C69. In addition, O6 of base G2 is held by the main chain NH group of Gly43. All these contacts at the extremity of the loop account for the ordering of this region in the complex, while loop 1 was disordered from residues 40 to 45 in the free enzyme crystals (Schmitt *et al.*, 1996a). Moreover, the interaction of Arg42 in the major groove of the acceptor helix may explain the results of a previous NMR analysis (Wallis *et al.*, 1995), where a general broadening and loss of intensity of resonances assigned to G–C base pairs in the $tRNA_f^{Met}$ acceptor stem were observed upon complex formation with the

Table I. Catalytic parameters of the formylase variants

	K_{cat} (s^{-1})	K_m (μM)	Relative K_{cat}/K_m (%)	Interactions with tRNA of the residue under study
WT	28 ± 2	0.2 ± 0.1	100	
R42K	2.6 ± 0.1	2.8 ± 0.3	0.7	O6 of G70 (Nε); N7 of G70 (Nη)
R42A	>0.3	>10	0.02	O6 of G70 (Nε); N7 of G70 (Nη)
Δ38–47	>0.2	>10	0.01	
K206A	4.0 ± 0.3	1.0 ± 0.3	2.9	none
K209A	16 ± 1	2.5 ± 0.2	4.6	OP of C71 (Nζ); OP of G72 (Nζ)
K291A	27 ± 1	1.7 ± 0.2	12	OP of C13 (Nζ); O2′ of G6 (Nζ)
K292A	21 ± 1	1.4 ± 0.2	11	none
R303A	21 ± 1	2.3 ± 0.3	6.5	OP of A14 (Nη)
R304A	22 ± 1	0.50 ± 0.05	33	O2′ of U24 (Nη)
N301A	30 ± 4	0.8 ± 0.1	27	O2 of U24 (Nδ); O2′ C25 (Nδ); O2′ of U24 (Oδ)

Relative K_{cat}/K_m are expressed as percentages of the value for the wild-type enzyme. K_m is that of tRNA. In all cases, it was verified that the K_m of FTHF was not significantly modified (data not shown).

formylase. These changes were not detected when the elongator $tRNA^{Met}$ was used.

The acceptor stem and the D-stem approach the C-terminal domain on their minor-groove sides. Several interactions involving the phosphate backbone and the riboses O2′ take place along an electropositive channel leading to the active site in the N-domain. This channel features Lys209, Lys246, Lys291, Arg303 and Arg304, all of which participate to the binding of the acceptor helix and of the D-stem. At the bottom of the acceptor stem, the N2 group of G5 is held by the main chain carbonyl of Gly290. Other base-specific interactions involve bases U24 and G12. Nδ of Asn301 makes a hydrogen bond with O2 of U24 and the main chain carbonyl group of Asn301 contacts N2 of G12. The specific interaction of Asn301 with base U24 sustains the report (Lee *et al.*, 1991) that substitution of base pair A11–U24, a characteristic of several eubacterial initiatior tRNAs, by C11–G24 reduced the efficiency of formylation by a factor of 8. In Table I, substitution of Asn301 with alanine is shown to increase four-fold the K_m of Met-$tRNA_f^{Met}$.

Binding of the 3′ end inside the catalytic site

The 3′-terminal adenosine and the attached formyl-methionyl group dip into the active-site cavity. While bases 72–75 are stacked on top of each other, base 76 is unstacked and its ribose group adopts a C2′-endo conformation. A large peak of positive density, possibly a water molecule or a magnesium ion, contacts O2 of cytosine 75 and is stacked on the 3′-terminal base (Figure 1). Such a stacking might be important in the formylation reaction by stabilizing the ring of base 76.

The side chain of Lys206 at the entrance of the active site must be displaced to allow the positioning of the terminal adenosine. Such a movement may contribute to the rotations of the linker and of the C-domain in response to the binding of tRNA (Figure 2). The close proximity of the Lys206 side chain and the tRNA 3′ terminal ribose had already been indicated by labeling experiments using periodate-oxidized $tRNA_f^{Met}$ (Gite and RajBhandary, 1997). Phosphate 76 is clamped through interactions involving the NH main chain group of Gly91 and the side chains of Thr11 and Gln35. Inside the active-site crevice, between strands β5 and β4 of the nucleotide binding fold, a strong density could be unambiguously attributed to the formyl-methionyl group esterified to the 3′-end adenosine. The side chain of methionine dips into a cavity surrounded by the hydrophobic side chains of Phe14, Ile123, Leu136, Leu171, Ala89, Pro122 and Tyr168 (Figure 1). This cavity appears well-suited for a methionine side chain. Its hydrophobic character also accounts for the capacity of formylase to react, although with reduced efficiency, with $tRNA_f^{Met}$ that has been misesterified by other amino acids (Giegé *et al.*, 1973; Li *et al.*, 1996). On the other hand, the formyl group is tightly held by the main chain NH groups of Ala89 and Tyr90, as well as by the side chain of Asn108. This asparagine is highly conserved in the family of FTHF-utilizing enzymes and is believed to be important in the catalytic mechanism of GARF from *E.coli* (Almassy *et al.*, 1992).

Upon superimposition of the structure of GARF complexed with the glycinamide ribonucleotide (GAR) substrate to that of the N-terminal domain of the complexed formylase, the position of the amino group of GAR, the formyl acceptor, appears to coincide almost exactly with that of the amino group of the methionine esterified to the tRNA. However, the ribose group in the GAR substrate does not superimpose to the terminal ribose of tRNA. Instead, it occupies the positions of the methionine side chain and of the Phe14 residue of the formylase.

Mechanism of tRNA discrimination

Upon its docking to the formylase, the body of the inititator $tRNA_f^{Met}$ establishes two sets of contacts. (i) The acceptor helix and the D-stem interact with the C-terminal domain of the enzyme in a rather unspecific manner. Contacts mainly involve the phosphate backbone and a cationic channel at the enzyme surface. Upon systematic mutation into alanine of cationic residues delineating the above channel, 2.5- to 10-fold increases in the K_m values of Met-$tRNA_f^{Met}$ are observed (Table I). K_{cat} values remain insensitive to the mutations, showing, therefore, that the free energy of binding ensuing from the C-domain is not used to improve the catalytic rate inside the N-domain. (ii) The second region of contact involves the top of the acceptor stem and loop 1. As shown in Figure 3, in order to dip into the active site crevice, the acceptor end of the bound tRNA must strongly deviate from the helical conformation generally encountered in the case of free or liganded tRNAs (Figure 5). Such a bending is reminiscent

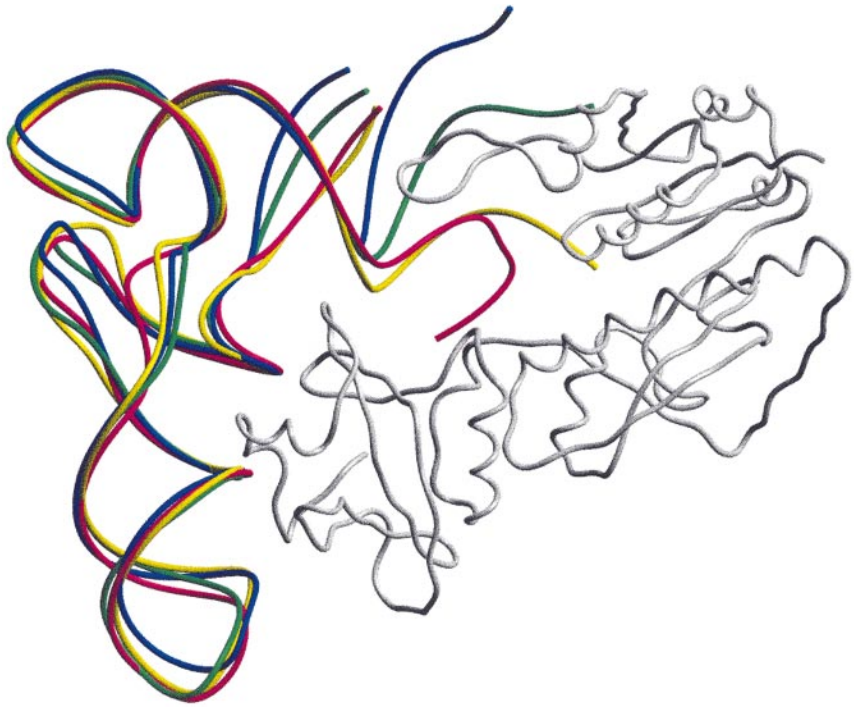

Fig. 5. Comparison of various tRNAs with the formylase-bound $tRNA_f^{Met}$. Free $tRNA^{Phe}$ (green), EF-Tu bound $tRNA^{Phe}$ (blue) and GlnRS bound $tRNA^{Gln}$ (pink) were superimposed to $tRNA_f^{Met}$ (yellow) complexed to the formylase (gray). Superimposition was based on the fitting of the phosphate atoms of tRNAs, excluding those of the acceptor arms. The C_α trace of the enzyme and the phosphate backbone traces of tRNAs were drawn using Setor (Evans, 1993).

of the case of $tRNA^{Gln}$ complexed with *E.coli* GlnRS, in which strong deviation from the helical conformation of the acceptor stem is also accompanied by disruption of the 1–72 pair (Rould *et al.*, 1989). The functional significance of the disruption of the 1–72 pair in $tRNA^{Gln}$ was reinforced by the observation that an amber $tRNA_f^{Met}$ (which is glutaminylatable because of the introduction of a U35 in the centre of the anticodon) loses aminoacylability by GlnRS if its C1–A72 mismatch is changed into the stable C1–G72 pair (Schulman and Pelka, 1985; Varshney *et al.*, 1991). In the case of the initiator tRNA–formylase complex, one consequence of the opening of the 1–72 bases is to allow loop 1 to establish contacts with base pairs 2–71 and 3–70, as well as with the phosphate backbone in the acceptor arm. As a result, the acceptor end is forced to dip into the active-site crevice (Figure 3). In this process, an A at position 73 appears optimal. In the case of a G at this position, a bad contact involving the N2 amino group of the base and the peptidic backbone of loop 1 would be created at the level of Ala40 and Gly41. All the above observations give a structural basis to the function of the known determinants for formylation clustered in the acceptor stem of initiator tRNA (Lee *et al.*, 1991; Guillon *et al.*, 1992b).

Construction of a formylase with residues 38–47 replaced by a Leu–Gly–Gly tripeptide (Table I) shows further the crucial role of loop 1 in determining the unique specificity of formylase. The formylating efficiency of the resulting enzyme (Δ38–47) is decreased by four orders of magnitude. In addition, the truncated formylase responds poorly now to the presence or absence of a strong base pair at position 1–72 in the substrate (Table II). Conversely, an R42A enzyme mutant discriminates much better the presence of a 1–72 mismatch in its substrate, most probably because the length of loop 1 has been left intact. On the other hand, the change of the side chain of Arg 42 into an alanine renders the K_m of met-$tRNA_f^{Met}$ immeasurably high. The introduction of a lysine at position 42 is not sufficient to restore the key interactions associated with the side chain of Arg42 (Table I). The dominant role of loop1 in the expression of enzyme specificity is also illustrated by Ramesh *et al.* (1997), by whom G41R and G41K mutant enzymes are reported to compensate for the strong negative effect accompanying the introduction of both A72G and A73G changes in an amber $tRNA_f^{Met}$.

Table II. Sensitivity to the 1–72 base pair in tRNA of three formylase variants altered in loop 1

	Wild type	Δ38–47	R42A
$tRNA_f^{Met}$ C1–A72 (wild-type)	100	100	100
$tRNA_f^{Met}$ C1–G72	0.04	3.3	0.6
$tRNA_f^{Met}$ G1–C72	0.01	13	1.0

Catalytic efficiencies (K_{cat}/K_m) towards the three indicated methionylated tRNAs were measured for wild-type, Δ38–47 and R42A enzymes. Values are expressed as percentages of those obtained with wild-type $tRNA_f^{Met}$ (see Table I).

Comparison with other tRNA-binding proteins

High specificity of the formylase towards its tRNA substrate is achieved mostly through recognition of the acceptor arm. This distinguishes the formylase from the family of aaRS in which the recognition always involves an additional region of the tRNA molecule, most often the anticodon loop (Delarue and Moras, 1993). The

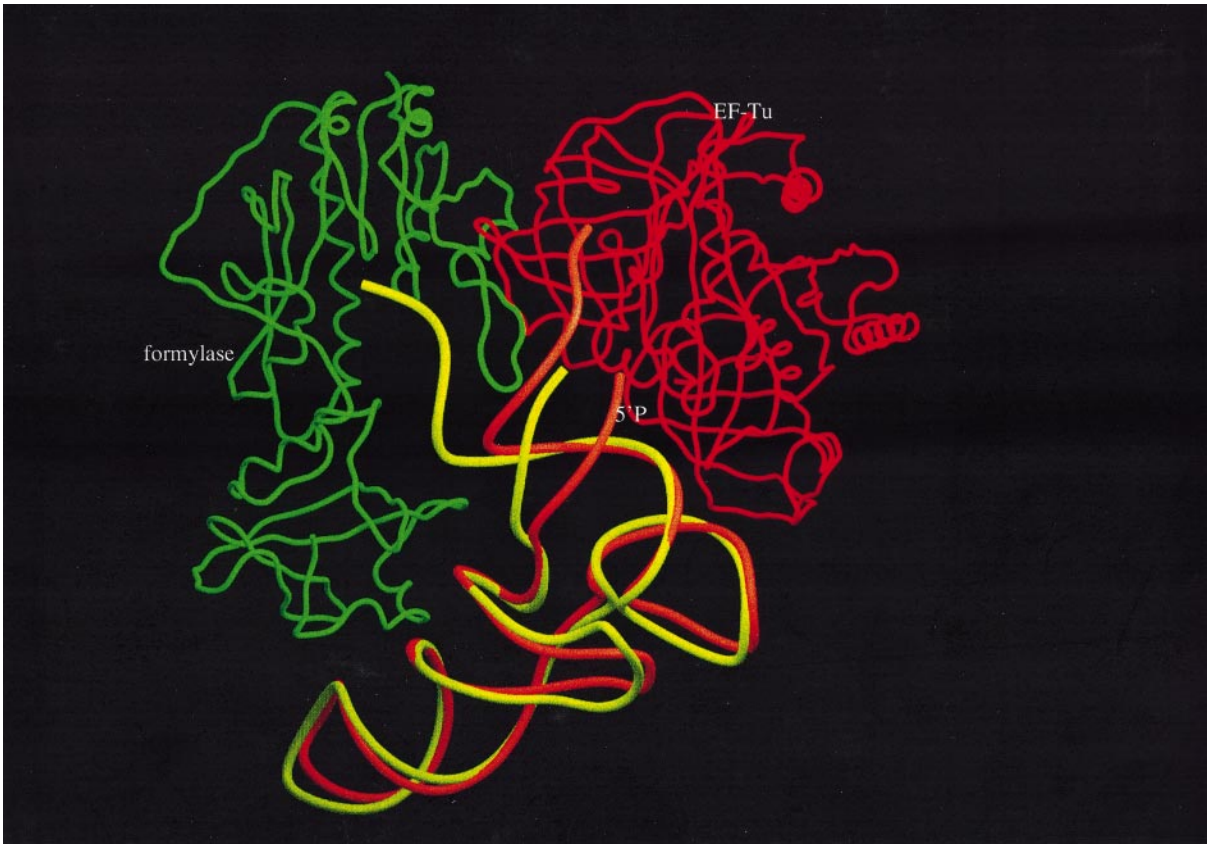

Fig. 6. Comparison of the binding of a tRNA by the formylase or by the elongation factor Tu. The EF-Tu–tRNA^{Phe} (red) (Nissen *et al.*, 1995) and formylase–tRNA$_f^{Met}$ (green) complexes were superimposed, according to a fit of the tRNA molecules which excludes the acceptor arms. The figure shows the different approaches of a tRNA ligand by the two protein molecules.

specificity of elongation factor Tu towards the set of elongator tRNAs is satisfied mainly through the binding of the acceptor arm, as in the case of formylase (Nissen *et al.*, 1995). However, tRNAs of these systems are recognized through opposite sides: EF-Tu interacts with tRNA on the T-stem side, whereas formylase approaches tRNA on the D-stem side (Figure 6). Moreover, the 5′ phosphoryl group of elongator tRNA is tightly bound at the surface of EF-Tu, whereas base 1 is away from the protein surface in the tRNA$_f^{Met}$–formylase complex. With the formylase, as well as with EF-Tu, bending of the acceptor end enables the esterified amino acid to reach its site. Nevertheless, the induced bendings have opposite directions.

In vivo studies indicate that slight overexpression of EF-Tu decreases the initiator activity of a tRNA by misappropriation (Guillon *et al.*, 1996). One may therefore suspect that, even under wild-type conditions, formylase undergoes the competition of EF-Tu for the uptake of an aminoacylated tRNA$_f^{Met}$. This view is supported by *in vitro* experiments showing that EF-Tu.GTP binds Met-tRNA$_f^{Met}$ with a dissociation constant in the 10 nM range, only 10-fold greater than the K_d of an elongator tRNA such as Met-tRNA$_m^{Met}$ (Janiak *et al.*, 1990). Such a competition may be partly overcome by the possibility illustrated in Figure 6 of simultaneously docking a formylase molecule and an elongation factor on the same tRNA. In the resulting complex, formylase could compete with the factor for the appropriation of the 3′ end of initiator tRNA, without the requirement for prior full dissociation of the tRNA molecule. Movement of the tRNA in the direction of the formylase might be favored by the C1–A72 mismatch and the ensuing abnormal mobility of the 5′ phosphate. Once formylation has succeeded, the tRNA irreversibly escapes association with EF-Tu and becomes committed to IF2 binding.

Materials and methods

Crystallization and data collection

Formylase and fMet-tRNA$_f^{Met}$ were purified as described previously (Schmitt *et al.*, 1998). Suitable crystals for X-ray experimentation were obtained using macroseeding techniques. Optimal conditions for growth correspond to 23% polyethylene glycol 5000 monomethylether, 0.2 M ammonium sulfate, 8% ethylene glycol, 60 μM fMet-tRNA$_f^{Met}$, 66 μM formylase, 10 mM magnesium chloride and 50 mM potassium chloride in 50 mM sodium morpholinoethanesulfonate buffer, pH 6.6 (Schmitt *et al.*, 1998). Crystals were orthorhombic, space group $P2_12_12$, with unit cell parameters of a = 201.7 Å, b = 68.1 Å, c = 86.4 Å. One mercury derivative was prepared by soaking crystals in 5 mM sodium parachloromercuriphenylsulfonate (PCMBS). Because of the presence of ethylene glycol, crystals could be directly cooled in liquid ethane for data collection. Data were collected at 100K by using a synchrotronic source (λ = 1.0 Å) at the LURE (Orsay, France) on a MAR-Research phosphor image plate system (Hamburg, Germany). Diffraction images were analyzed using the MOSFLM program (A.G.W.Leslie, Laboratory of Molecular Biology, Daresbury, UK) and the data further processed using programs from the CCP4 package (Collaborative Computational Project No. 4, 1994).

Phase determination and model building

Calculation of a native Patterson map (12–4.0 Å resolution) showed a large positive peak reaching 18% of the origin peak at fractional

Table III. Data collection and refinement statistics

Data	Native	PCMBS
Resolution used (Å)	30–2.8	30–2.92
Completeness (%)	99.4	94.8
Redundancy	3.3	3.0
R_{sym} (I)[a]	5.1 (35)	6.5 (35)
Mean intensity[b]	1931 (1082;618;330;182)	2533 (1422;886;482;258)
ΔF_{iso}[c] (%)		23.9
Heavy atom sites		4
R_{cullis}[d]		0.67
Refinement		
Reflections used in refinement	24 608	
R-factor (%)	24.7	
Free-*R*-factor[e] (%)	29.2	
r.m.s.d. bonds (Å)	0.01	
r.m.s.d. angles (deg.)	1.4	

Each data set was collected with a single crystal. The derivative crystal was soaked in 5 mM sodium parachloromercuriphenylsulfonate (PCMBS) for 4 days. Because of the presence of ethylene glycol, crystals could be directly cooled in liquid ethane for data collection. Data were collected at 100K by using a synchrotron source (λ = 1.0 Å) at the LURE (Orsay, France) on a MAR-Research phosphor image plate system (Hamburg, Germany). Diffraction images were analyzed with the MOSFLM program (A.G.W.Leslie, Laboratory of Molecular Biology, Daresbury, UK) and the data further processed using programs from the CCP4 package (Collaborative Computational Project No. 4, 1994).

[a]R_{sym} (I) = $\frac{\Sigma_{hkl}\Sigma_i|(I_{hkl}) - I_{hkl,i}}{\Sigma_{hkl}\Sigma_i|I_{hkl}}$ where *i* is the number of reflections *hkl*. The value in parentheses corresponds to the highest resolution shell.

[b]The overall value is indicated. Values in parentheses correspond to the four highest resolution shells (shell limits are 3.61, 3.35, 3.13, 2.95 and 2.80 Å for native data, and 3.76, 3.49, 3.26, 3.08 and 2.92 Å for derivative data).

[c]$\Delta F_{iso} = \frac{\sum_{hkl}|Fph_{hkl} - Fp_{hkl}|}{\sum_{hkl}Fp_{hkl}}$ where Fph and Fp are the structure factors of the derivative and of the native crystal, respectively.

[d]$R_{cullis} = \frac{\sum_{hkl}|Fph - \vec{Fp} + \vec{Fh}||}{\sum_{hkl}|Fph - Fp|}$ for the centric terms only, where Fh are the structure factors of the heavy atom.

[e]Calculated for 1528 reflections left out of the refinement procedure.

coordinates (0.0, 0.5, 0.0). This strongly suggested the occurrence of two molecules of complex related by an almost pure NCS translation of a half-cell edge along the *b*-axis. In addition, the positions of the mercury atoms of the PCMBS derivative were identified by analysis of the Patterson difference map. The four mercury sites clearly divided into two groups related to one another by the expected NCS operator. The positions were refined and used to calculate SIR (single isomorphous replacement) phases and moduli using the SHARP program (de La Fortelle and Bricogne, 1997) combined with solvent flattening using Solomon (Collaborative computational project No. 4, 1994). Maps were calculated at 4.0, 3.5 and 3.0 Å resolution. Three pieces of evidence gave identical positions for the formylase molecules: (i) two copies of the native formylase enzyme could be unambiguously fitted in the SIR electron density; (ii) in this positioning the mercury atoms are bound to C154 and C226, which are the main mercury sites in the native enzyme crystal (Schmitt *et al.*, 1996a); (iii) the fitting into the SIR map is in agreement with the best solution obtained by molecular replacement (AMoRe; Navaza, 1994). Finally, after the two enzyme molecules had been positioned, large parts of remaining density could clearly be attributed to two tRNA molecules. The known 3D-structures of tRNAs were used as guides to construct a starting model containing all stems and the acceptor arm.

Refinement

The starting model corresponding to one of the two NCS-related complexes was manually fitted into the SIR density to obtain the second complex. The positions of the four molecules were then refined as rigid bodies. After this step, the *R*-factor was 45% (free-*R*, calculated with 6% of the data, was 47%). The final positions allowed us to to refine slightly different NCS operators for the tRNA molecules and for the enzyme molecules. Therefore, refinement was performed by maintaining tight NCS restraints split into two groups containing the protein molecules and the tRNAs, respectively, excluding regions involved in crystal packing. The model was refined by applying cycles of torsion angle molecular dynamics and positional refinement with X-PLOR 3.851 (Brunger, 1992; Rice and Brunger, 1994) and later with CNS (Pannu and Read, 1996; Adams *et al.*, 1997), alternated with manual rebuilding using program O (Jones *et al.*, 1991). During this process, the construction of the whole tRNA molecule could be achieved, as well as the construction of all enzyme residues including residues 39–45, which were absent from the starting model. The final model comprises the two complexes, 83 water molecules and two magnesium ions, and has good geometry, with r.m.s. deviations from ideal values of 0.01 Å for bond lengths and 1.4° for bond angles. Final *R* and free-*R*-factors are, respectively, 24.7% for 24 608 reflections (two sigmas cut-off), and 29.2% for 1528 reflections left out of the refinement. One complex is on the whole slightly more disordered than the other. However, neither the free-*R*-factor nor the quality of the density could be improved by refining the structure without NCS restraints. In the tRNA molecules, the D loop (residues 16–19), as well as bases 1 and 37, appeared more disordered, even though the positions of these bases could be tentatively modeled. The average B factors are 62.7 $Å^2$ (overall), 48 $Å^2$ (formylase molecule 1), 75 $Å^2$ (tRNA molecule1), 59 $Å^2$ (formylase molecule 2) and 77 $Å^2$ (tRNA molecule 2). These high values are consistent with the rapid decay of the reflection intensities below 3.5 Å resolution (Table III). Despite this, the final density is generally of good quality (Figure 1).

Production and characterization of mutant enzymes

Two fragments (*Xba*I–*Nco*I and *Nco*I–*Bam*HI) of the *fmt* gene from pUC18Fatg (Schmitt *et al.*, 1996c) were subcloned into a modifed M13mp18 phage vector. Oligonucleotide site-directed mutagenesis was then performed by using the resulting single-stranded DNAs as template. The mutated genes were re-inserted into the pUC18Fatg expression vector and completely sequenced. The corresponding enzymes were produced free of contamination by wild-type formylase in the *fmt* null strain PAL13Tr (Guillon *et al.*, 1992a), and purified to homogeneity as described (Schmitt *et al.*, 1996c).

Initial rates of Met-$tRNA_f^{Met}$ formylation in the presence of catalytic

amounts of the studied enzymes were measured as described previously (Blanquet *et al.*, 1984; Guillon *et al.*, 1992b) in a buffer (20 mM Tris pH 7.6, 0.1 mM EDTA, 10 mM 2-mercaptoethanol, 150 mM KCl, 7 mM $MgCl_2$) containing 125 μM FTHF and 0.05–15 μM Met-$tRNA_f^{Met}$. $tRNA_f^{Met}$ and its derivatives used in the assay were overproduced *in vivo* from synthetic genes and purified as described previously (Meinnel *et al.*, 1988; Meinnel and Blanquet, 1995) to methionine acceptances ranging between 1500 and 1700 picomoles per A_{260} unit.

Acknowledgements

The authors gratefully acknowledge Luc Moulinier for helpful advice, Pierre Plateau and Frederic Dardel for critical reading of the manuscript, and the staff of the LURE DW32 beamline for assistance during data collection. Coordinates have been deposited at the Brookhaven Protein Data Bank (accession No. 2fmt).

References

Adams,P.D., Pannu,N.S., Read,R.J. and Brunger,A.T. (1997) Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. *Proc. Natl Acad. Sci. USA*, **94**, 5018–5023.

Almassy,R.J., Janson,C.A., Kan,C. and Hostomska,Z. (1992) Structures of apo and complexed *Escherichia coli* glycinamide ribonucleotide transformylase. *Proc. Natl Acad. Sci. USA*, **89**, 6114–6118.

Blanquet,S., Dessen,P. and Kahn,D. (1984) Properties and specificity of methionyl-$tRNA_f^{Met}$ formyltransferase from *Escherichia coli*. *Methods Enzymol.*, **106**, 141–152.

Brunger,A.T. (1992) Free-R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–474.

Chen,P., Schulze-Gahmen,U., Stura,E.A., Inglese,J., Johnson,L.D., Marolewski,A., Benkovic,S.J. and Wilson,I.A. (1992) Crystal structure of glycinamide ribonucleotide transformylase from *Escherichia coli* at 3.0 Å resolution. *J. Mol. Biol.*, **227**, 283–292.

Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.

de La Fortelle,E. and Bricogne,G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.*, **276**, 472–494.

Delarue,M. and Moras,D. (1993) The aminoacyl-tRNA synthetase family: modules at work. *Bioessays*, **15**, 675–687.

Evans,S.V. (1993) Setor: hardware lighted three-dimensional solid model representation of macromolecules. *J. Mol. Graphics*, **11**, 134–138.

Giegé,R., Ebel,J.P. and Clark,B.F.C. (1973) Formylation of mischarged *E.coli* $tRNA_f^{Met}$. *FEBS Lett.*, **30**, 291–295.

Gite,S. and RajBhandary,U.L. (1997) Lysine 207 as the site of crosslinking between the 3′ end of *Escherichia coli* initiator tRNA and methionyl-tRNA formyltransferase. *J. Biol. Chem.*, **272**, 5305–5312.

Guillon,J.M., Mechulam,Y., Schmitter,J.M., Blanquet,S. and Fayat,G. (1992a) Disruption of the gene for Met-$tRNA_f^{Met}$ formyltransferase severely impairs growth of *Escherichia coli*. *J. Bacteriol.*, **174**, 4294–4301.

Guillon,J.M., Meinnel,T., Mechulam,Y., Lazennec,C., Blanquet,S. and Fayat,S. (1992b) Nucleotides of tRNA governing the specificity of *Escherichia coli* methionyl-$tRNA_f^{Met}$ formyltransferase. *J. Mol. Biol.*, **224**, 359–367.

Guillon,J.M., Heiss,S., Suturina,J., Mechulam,Y., Laalami,S., Grunberg-Manago,M. and Blanquet,S. (1996) Interplay of methionine tRNAs with translation elongation factor Tu and translation initiation factor 2 in *Escherichia coli*. *J. Biol. Chem.*, **271**, 22321–22325.

Janiak,F., Dell,V.A., Abrahamson,J.K., Watson,B.S., Miller,D.L. and Johnson,A.E. (1990) Fluorescence characterization of the interaction of various transfer RNA species with elongation factor Tu.GTP: evidence of a new functional role for elongation factor Tu in protein biosynthesis. *Biochemistry*, **29**, 4268–4277.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for the building of proteins model in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.

Kahn,D., Fromant,M., Fayat,G., Dessen,P. and Blanquet,S. (1980) Methionyl-transfer-RNA transformylase from *Escherichia coli*: purification and characterisation. *Eur. J. Biochem.*, **105**, 489–497.

Ladner,J.E., Jack,A., Robertus,J.D., Brown,R.S., Rhodes,D., Clark,B.F. and Klug,A. (1975) Structure of yeast phenylalanine transfer RNA at 2.5 Å resolution. *Proc. Natl Acad. Sci. USA*, **72**, 4414–4418.

Lee,C.P., Seong,B.L. and RajBhandary,U.L. (1991) Structural and sequence elements important for recognition of *Escherichia coli* formylmethionine tRNA by methionyl-tRNA transformylase are clustered in the acceptor stem. *J. Biol. Chem.*, **266**, 18012–18017.

Li,S., Kumar,N.V., Varsney,U. and RajBhandary,U.L. (1996) Important role of the amino acid attached to tRNA in formylation and in initiation of protein synthesis in *Escherichia coli*. *J. Biol. Chem.*, **271**, 1022–1028.

Mandal,N., Mangroo,D., Dalluge,J.J., Mac Closkey,J.A. and RajBhandary,U.L. (1996) Role of the three consecutive G:C base pairs conserved in the anticodon stem of initiator tRNAs in initiation of protein synthesis in *Escherichia coli*. *RNA*, **2**, 473–482.

Meinnel,T. and Blanquet,S. (1995) Maturation of pre-$tRNA^{fMet}$ by *E.coli* RNase P is specified by a guanosine of the 5′ flanking sequence. *J. Biol. Chem.*, **270**, 15906–15914.

Meinnel,T., Mechulam,Y. and Fayat,G. (1988) Fast purification of a functional elongator $tRNA^{Met}$ expressed from a synthetic gene *in vivo*. *Nucleic Acids Res.*, **16**, 8095–8096.

Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.

Nissen,P., Kjeldgaard,M., Thirup,S., Polekhina,G., Reshetnikova,L., Clark,B.F.C. and Nyborg,J. (1995) Crystal structure of the ternary complex of Phe-$tRNA^{Phe}$, EF-Tu and a GTP analog. *Science*, **270**, 1464–1472.

Pannu,N.S. and Read,R.J. (1996) Improved structure refinement through maximum likelihood. *Acta Crystallogr.*, **A52**, 659–668.

Ramesh,V., Gite,S., Li,Y. and RajBhandary,U.L. (1997) Suppressor mutations in *Escherichia coli* methionyl-tRNA formyltransferase: role of a 16-amino acid insertion module in initiator tRNA recognition. *Proc. Natl Acad. Sci. USA*, **94**, 13524–13529.

Rice,L.M. and Brunger,A.T. (1994) Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins*, **19**, 277–290.

Rould,M.A., Perona,J.J., Söll,D. and Steitz,T.A. (1989) Structure of *E.coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8 Å resolution. *Science*, **246**, 1135–1142.

Schmitt,E., Blanquet,S. and Mechulam,Y. (1996a) Structure of crystalline *Escherichia coli* methionyl-$tRNA_f^{Met}$ formyltransferase: comparison with glycinamide ribonucleotide formyltransferase. *EMBO J.*, **15**, 4749–4758.

Schmitt,E., Guillon,J.M., Meinnel,T., Mechulam,Y., Dardel,F. and Blanquet,S. (1996b) Molecular recognition governing the initiation of translation in *Escherichia coli*. A review. *Biochimie*, **78**, 543–554.

Schmitt,E., Mechulam,Y., Ruff,M., Mitschler,A., Moras,D. and Blanquet,S. (1996c) Crystallization and preliminary X-Ray analysis of *Escherichia coli* methionyl-$tRNA_f^{Met}$ formyltransferase. *Proteins*, **25**, 139–141.

Schmitt,E., Blanquet,S. and Mechulam,Y. (1998) Crystallization and preliminary X-Ray analysis of *Escherichia coli* methionyl-$tRNA_f^{Met}$ formyltransferase complexed with formyl-methionyl-$tRNA_f^{Met}$. *Acta Crystallogr.*, **D**, in press.

Schulman,L.H. and Pelka,H. (1985) *In vitro* conversion of a methionine to a glutamine-acceptor tRNA. *Biochemistry*, **24**, 7309–7314.

Suddath,F.L., Quigley,G.J., McPherson,A., Sneden,D., Kim,J.J., Kim,S.H. and Rich,A. (1974) Three-dimensional structure of Yeast phenylalanine transfer RNA at 3.0 Å resolution. *Nature*, **248**, 20–24.

Varshney,U., Lee,C.P. and RajBhandary,U.L. (1991) Direct analysis of aminoacylation levels of tRNAs *in vivo*. *J. Biol. Chem.*, **266**, 24712–24718.

Wallis,N.G., Dardel,F. and Blanquet,S. (1995) Heteronuclear NMR studies of the interactions of the ^{15}N-labeled methionine-specific transfer RNAs with methionyl-tRNA transformylase. *Biochemistry*, **34**, 7668–7677.

Woo,N.H., Roe,B.A. and Rich,A. (1980) Three-dimensional structure of *Escherichia coli* initiator $tRNA_f^{Met}$. *Nature*, **286**, 346–351.

Received September 3, 1998; revised and accepted September 28, 1998